

Client-Server, Distributed Database Strategies in a Healthcare Record System for a Homeless Population

Henry C. Chueh, M.D., M.S., and G. Octo Barnett, M.D.

Laboratory of Computer Science, Massachusetts General Hospital

A computer-based healthcare record system being developed for Boston's Healthcare for the Homeless Program (BHCHP) uses client-server and distributed database technologies to enhance the delivery of healthcare to patients of this unusual population. The needs of physicians, nurses and social workers are specifically addressed in the application interface so that an integrated approach to healthcare for this population can be facilitated. These patients and their providers have unique medical information needs that are supported by both database and applications technology. To integrate the information capabilities with the actual practice of providers of care to the homeless, this computer-based record system is designed for remote and portable use over regular phone lines. An initial standalone system is being used at one major BHCHP site of care. This project describes methods for creating a secure, accessible, and scalable computer-based medical record using client-server, distributed database design.

INTRODUCTION

A computer-based healthcare record system developed for Boston's Healthcare for the Homeless Program (BHCHP) helps providers manage the diverse healthcare needs of Boston's homeless population. Recent studies of the homeless population have emphasized the complexity of healthcare issues in this group [9]. Not only do these patients have complex healthcare issues, they also obtain their health services in non-traditional ways. For example, routine patient care by BHCHP occurs at multiple sites with different types of providers at different sites, making access to site-based paper medical records nearly impossible. The transient nature of this population can make access to the medical record critical. Increasing requirements for comprehensive documentation of health care delivery and the growing number and diversity of the homeless population also creates a challenge for providers managing these patients.

We are meeting these needs with client-server, distributed database technologies and attention to a

flexible applications interface for providers that is patient-oriented in terms of both accessibility and design. An initial standalone system has been installed at the BHCHP Respite Care Unit to capture registration data and to prototype the interface to the database. Information collected through the daily use of this computer-based record will enhance the ability to analyze the prevalence of important disease entities within the population, and to help assess which treatment options are successful in this group. The more extensive client-server, remote solution is currently under active development with major components of the database schema, user interface, and distributed database strategies completed. This project implements methods and strategies which can increase the availability of healthcare information beyond traditional care sites and promote less provider- or site-oriented care and more patient-oriented care.

THE HOMELESS POPULATION

Patterns of care

Homeless patients are often seen by different BHCHP providers at multiple sites in the city. These sites include homeless shelters, major hospitals, a BHCHP-operated respite unit, and a mobile van. While acute hospitalizations for these patients may not differ significantly from most, the outpatient care of the homeless population is radically different. Patients are seen frequently on an ad hoc basis for symptomatic complaints without regard to a primary site or provider of care.

Furthermore, the disposition of homeless patients is particularly complex, since a patient who is well enough to return home may not be well enough to return to the streets or a shelter. While some patients do have a primary care provider and site of care, most of the population remains disconnected from traditional health care access [1]. Follow-up care is also particularly difficult to arrange, and will often occur extemporaneously with a different provider at another site. Although the average ambulatory care patient may also have multiple care providers through referrals, a primary care provider usually coordinates the care. The multiplicity of time, place and person in these patients' routine

health care delivery makes traditional practice-based records all but unusable.

Healthcare management needs

Besides unusual patterns of health care delivery, BHCHP's population has unique needs for medical record information because of the prevalence of specific disease entities in the homeless population coupled with the frequent lack of follow-up care [1, 4]. Some of these problems such as substance abuse, AIDS, and tuberculosis, can pose significant health hazard to the general population if they are endemic in the homeless population. However, because of the current difficulty in obtaining medical record information about PPD test results, chest x-rays, and anti-tuberculosis treatments, managing these types of problems in this high risk population is increasingly difficult [3].

Often the critical issue in overall healthcare for these patients is not a specific medical problem, but the management of their social welfare [2]. Solving these problems requires a coordinated effort of care between medical and social providers [9]. Integrated efforts can aid these patients in obtaining temporary shelter, Medicaid insurance, low-cost housing, social security, Veterans' benefits and other benefits of state programs. A computer-based healthcare system for these patients must integrate these critical tasks and their associated information demands.

DATABASE STRATEGY

Our solution for access to medical records from multiple sites by multiple providers is a client-server architecture. Local, client-based database capabilities are also implemented to enhance performance and provide a more interactive presentation than traditional terminal-based solutions.

Central database server

A transaction-based, relational database server maintains all patient medical record information. In addition to being a central repository for patient data, the server also handles security issues by providing and supporting techniques that grant access to the data, such as passwords and callback numbers. As a transaction-based database, the server responds only to database queries and updates, and does not participate in the applications interface at all. We are using Oracle 7 database software running as an NLM on a Novell network server.

Local client database

The local databases reside on the client computers which are fixed workstations at a designated sites or notebook computers for portable use. The local database is also implemented using relational database software. Initially we used IBM PC-based FoxPro/DOS software, but for consistency and transparency (FoxPro does not support a fully ANSI-compliant SQL as does Oracle), we are now using WATCOM SQL for Windows. These local relational tables provide several capabilities; first, they support temporary "working" storage for records retrieved from the central server; second, they allow the local storage of large, static vocabularies used to help control data entry by providers. For instance, a large diagnosis/problem vocabulary consisting of over 5000 entries can be used to validate entries into a problem list in 1 - 2 seconds. Other vocabularies include a full USP drug vocabulary, and a vocabulary of surgical procedures.

While vocabularies can be located on the central server, our tests show that vocabulary queries against the server are several times slower if the user-selected diagnostic term is not an exact match with a vocabulary item, and an interactive display of possible matches is provided. This delay occurs because of the data passing over the low-bandwidth telephone link (in our system architecture), and not because of the client-server model-- tests by us and others using the server-based vocabularies on a local area network achieved performance similar to local vocabularies [7]. In our case, local vocabularies provide rapid response even with broad queries which result in a large number of choices, and allow a provider to pick an entry from this list, rather than reformulating a more narrow query. An example of such a query is the entry of the term: "infection", which returns over 100 entries from the vocabulary.

CONNECTIVITY

Hardware

For remote systems, connecting the client computer with the server is a central issue. Given the high cost and difficulty in using dedicated wide-area networks and the infancy of wireless communications, we use regular phone lines with high-speed modems to connect remote sites to the server. Both modems and regular phone lines are low-cost and tested technologies. Unfortunately, this type of connection is relatively low in bandwidth, meaning that transfer of data must be minimized to achieve respectable performance. The client-server architecture is ideal for minimizing network traffic.

Multi-user, concurrent access to the database is handled by a network operating system. The relational database software runs on a dedicated database server computer that also runs a local area network operating system. This scheme allows any node on the network to be a client -- remote users are simply attached to the network through "network modems" (Shiva Corp. NetModem and LanRover) which negotiate the modem to network translation [see Fig. 1].

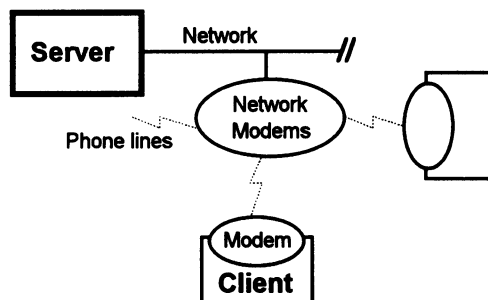


Figure 1. Layout of the client-server architecture using remote telephone links.

Software

Another way we are achieving reasonable real-time performance is to use Structured Query Language (SQL) to request information from the central server. SQL is a non-procedural query language based on relational calculus which is used by many relational databases as a standard query language [12]. The main advantage of SQL is the ability to query a database without procedural code. Only an SQL text string is sent to the database server, and the database server software will decide the optimal way to satisfy the query. In the specific case of our low-bandwidth client-server link, using SQL will limit the traffic over the link -- SQL queries to the server, and query results in return. Using transaction-based servers can, however, place increased requirements on the client software [6].

DATABASE SECURITY

Physical access

Secured access to the central database from a remote site is achieved through a validated callback mechanism. When a remote user dials into the system, the user provides a user name, personal password, and site password and then disconnects. The user is validated from private database tables; if authorized, the central server performs a callback to the corresponding phone number to establish a connection. This helps to protect the system against use from unauthorized phone numbers, which can be

a major problem for dial-up systems. Maintaining a single central repository for all patient-based data minimizes physical security issues.

Record retrieval and integrity

Once a secure connection is established with the central database server, a provider identifies a patient record to retrieve. Retrieval of a single record is achieved by identifying information such as name, age, sex, and birthdate. For this population, we provide additional identification such as pseudonyms, identifying marks, and general appearance characteristics such as eye color. Retrieval of a record constitutes bringing over all the latest information such as the current problem list and medications, as well as a small amount of historical information such as the last several encounter notes. When a record is identified and retrieved, it is marked as "checked out" on the central database server. This allows the database server to prevent simultaneous updates to parts of the record which are under the control of one provider at a time, such as the management of the problem list. Other transactions such as posting lab results or entering a new note continue to be valid. This coordination of simultaneous database access will be discussed in detail below.

CLIENT APPLICATION INTERFACE

The software interface used by providers is created for ease of use and hides the underlying connection to the central database. Once the patient record is retrieved, all information in the healthcare record is managed interactively using the local application and database. Repeated server access is only required to retrieve additional historical information, such as old visit notes. This allows providers to browse and edit items without delay. The record is divided into logical sections: Demographics, Visit notes, Lab Results and Electronic Mail, and Summary. Our initial applications interface was completed in FoxPro, but we have since migrated to graphical front-end tools for Windows on PC-based systems.

The Summary section is the core of the system and consists of flexible problem and therapy lists maintained by the providers. These lists are maintained over time, so that history "threads" of particular problems are created [5]. Detailed vocabularies maintained in the local database help validate terms that can be entered into the lists, while also maintaining maximum flexibility of data

entry. To enter a new term onto the problem list, a provider types a phrase of one or more words or parts of words. This phrase is then decomposed into single words, translated through stop word, acronym, and synonym vocabularies, and then matched to final terms. The list of problem terms is based on COSTAR vocabulary[10], and consists of over 5000 entries. The therapy vocabulary is a combination of COSTAR terms and a full USP drug index.

A typical search on a phrase will return a list of possible terms in less than 2 seconds. Moreover, terms are ranked by how closely they match the original phrase. This list can be reduced or expanded without performing further queries against the entire vocabulary. Figure 2. shows the results of entering "ACUTE UTI" as a new problem. This type of processing, where the user identifies a few key words to select from thousands of terms in a vocabulary provides great flexibility when using a controlled vocabulary [7].

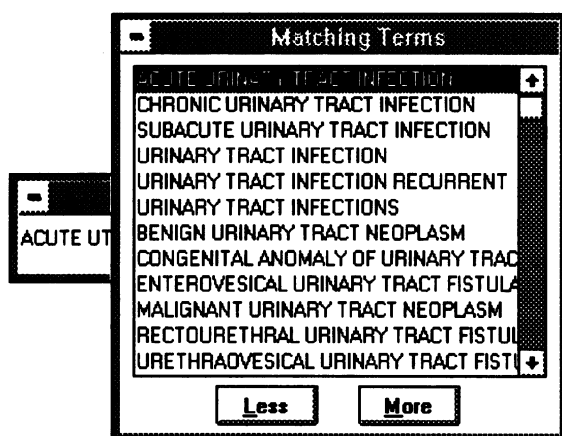


Figure 2. Results of expanding and matching the phrase "ACUTE UTI" against the local vocabulary.

Using these word-indexed vocabulary tables to support real-time matching allows us to include many actual, descriptive problems seen in this population. Problems such as "Recently arrested", "Recovering alcoholic", "Tuberculosis medication non-compliance" are difficult to categorize diagnostically -- but are of great importance to the overall welfare of these patients. Since we do not mandate a specific categorization of problems and we do allow providers to search for terms directly at the word level, these types of descriptive problems can be included usefully in the vocabulary.

COORDINATION OF DATABASES

In this client-server, distributed database environment, two kinds of data must be coordinated between the central host server and the remote clients: static and dynamic. Static data consists of databases such as the controlled vocabularies for problems and therapies. These can be maintained by issuing a maintenance query from the client to update the vocabularies on a periodic basis. In general, a "delta" update can be performed so that only new vocabulary items are added, and old ones removed, so that updates occur relatively quickly.

Dynamic data such as medical record information can require concurrent updates to maintain database integrity and to be useful to other users of the system. Therefore, this data is handled in a transaction-based fashion, with new data committed to the database once a transaction is complete. We have chosen to implement our system to handle both "short" and "long" transactions: Short transactions are those transactions which are committed to server in real-time to avoid inconsistencies in the medical record [8]. Data which is likely to have immediate use to concurrent users are treated as short transactions; for instance, the entry of lab results or e-mail messages. In our client-server architecture, there is no provision for automatically posting new results on the server to all connected clients. Initially, the distribution of short transactions will occur through a simple polling mechanism. The connected client program will periodically inspect a data field on the server (using a simple query with minimal overhead) which will indicate the presence of new short transaction information since logging on. Depending on the priority of the data, the data will either be retrieved automatically and the user will be alerted, or the system will ask the user if the data should be retrieved. We are exploring the possibility of using database triggers and message-passing programs to alert clients without requiring polling. On the other hand, a patient encounter is treated as a long transaction, since partial information about an encounter is not likely to be useful. Only after all encounter data has been entered (changes in the problem/therapy lists, new visit note, new lab test orders, etc.) does the encounter transaction get posted to the central server.

ANALYSIS AND REPORTING

Initial reports will help BHCHP providers identify active patients for discussion at their monthly "rounds". Reports will also aid BHCHP providers in

providing documentation for reimbursement and ongoing grant-based funding. Currently these types of reports are performed by hand. When the system has captured a significant amount of data, other reports (which are currently almost impossible to create) can help to describe features of this relatively unknown population -- prevalence of specific diseases, common medical problems, provider visits annually, success rates of treatments, and so on.

In addition to static reports, we have implemented "alert" features which provide some level of quality assurance to providers. Some of these alerts take the form of reminders; for instance, if a provider changes the status of a patient's PPD from negative to positive, and no chest x-ray is found in the record, the system could remind the provider to obtain a chest x-ray (conversely, if a recent x-ray had already been performed, it could remind the provider NOT to order another one). Other alerts will be intended to enhance communication between providers; for example, if one provider indicates in the record that a throat culture for strep was ordered, the next provider who retrieved that patient's record would see an alert to check the result of the culture. Such quality assurance reminders have proved useful [11].

CONCLUSION

We have used client-server, distributed database strategies to help develop a computer-based healthcare record system accessible from remote sites. The methods described above -- use of a transaction-based central database server, local vocabularies, and network/modem connectivity -- allow us to provide remote access while maintaining the integrity and security of the data stored. While this project meets the unusual demands of a homeless population, we expect that similar strategies will become increasingly important for developing a more patient-based care record for the general population.

An initial implementation of our system is operating as a stand-alone system at the BHCHP Respite Care Unit in Greater Boston. This patient-oriented record for a homeless population will provide the potential to identify more clearly the health care needs and problems of this population. Enhancing health care delivery to these people has been a difficult task not only in Boston, but also in other urban centers around the nation. Although this system is being

designed to support the needs of the BHCHP group and its patient population, we feel that the techniques involved are general. If the demand exists, dissemination of the program to help create a nationwide network of support for these patients and their providers would be extremely appealing. Such an expanded effort would require a practical and exciting collaboration between our Lab of Computer Science, BHCHP and the additional care sites.

Acknowledgment

Dr. Chueh is supported by NLM training grant 1 T15 LM07092. The authors also gratefully acknowledge support for this research by BHCHP and Hewlett-Packard Corp.

Reference

1. L. Gelberg, L.S. Linn, R.P. Usatine, M.H. Smith. Health, homelessness, and poverty. A study of clinic users. *Archives of Internal Medicine*. [JC:7fs] 150(11):2325-30, 1990 Nov.
2. M.A. Winkleby, B. Rockhill, D. Jatulis, S.P. Fortmann. The medical origins of homelessness. *American Journal of Public Health*. [JC:3xw] 82(10):1394-8, 1992 Oct.
3. K. Brudney, J. Dobkin. Resurgent tuberculosis in New York City. Human immunodeficiency virus, homelessness, and the decline of tuberculosis control programs. *American Review of Respiratory Disease*. [JC:426] 144(4):745-9, 1991 Oct.
4. D. McCarty, M. Argeriou, R.B. Huebner, B. Lubran. Alcoholism, drug abuse, and the homeless. Massachusetts Department of Public Health, Boston. *American Psychologist*. [JC:41v] 46(11):1139-48, 1991 Nov.
5. W. Beeler Jr., P.S. Gibbons, C.G. Chute. Development of a Clinical Data Architecture. *Proceedings of the Sixteenth Annual Symposium on Computer Applications in Medical Care*. 244-248, 1992.
6. S. Miller, C. Niedner, J. London. The Organization Engine: Virtual Data Integration. *Proceedings of the Sixteenth Annual Symposium on Computer Applications in Medical Care*. 610-614:1992.

7. T. Huff, T. A. Pryor, R. D. Tebbs. Pick From Thousands: A Collaborative Processing Model for Coded Data Entry. *Proceedings of the Sixteenth Annual Symposium on Computer Applications in Medical Care*. 104-108:1992.
8. A. Silberschatz, M. Stonebraker, J. Ullman. Database Systems: Achievements and Opportunities. *Communications of the ACM*. 34(10):110-120. 1991 Oct.
9. D. Stephens, E. Dennis, M. Toomer, J. Holloway. The diversity of case management needs for the care of homeless persons. *Public Health Reports - Hyattsville*. [JC:qja] 106(1):15-9, 1991 Jan-Feb.
10. G.O. Barnett. The application of computer-based medical record systems in ambulatory practice. *New Eng J Med*. 1984;310:1643-50.
11. G.O. Barnett, R.N. Winickoff, J.L. Dorsey, et. al. Quality assurance through automated monitoring and concurrent feedback using a computer-based medical information system. *Med Care*. 1978;16:962-970.
12. G. Wiederhold. Database Design. McGraw-Hill Co., 1983.